Understaing role of choices in marketing psychology using recommender systems

anand, atul

14 November 2019

Ingredient mix with recommender systems

Understaing role of choices in marketing psychology using recommender systems

Atul Anand
Bangalore, India
atulanandsh@gmail.com

***Abstract*—The product owners and marketers are always looking to launch new products and are presented constantly with the problem of what products or the variants they should present to the consumer. We present modern aspects of studies in this regard and combine them. We look deeper in the problems pertaining to earlier researches on product choices. We also provide a framework to the solution of this problem using recommender systems.**

Keywords: marketing psychology; malcolm gladwell; dan ariely; choice; consumer happiness, barry swartz, product design;

I. INTRODUCTION

The product owners and marketers have always looked to identify new ways to hook consumers to their product or their brand. On one hand they are looking for ways to gain new customers and on the other hand they do not want to lose “their” existing customers. They lose customers for multiple reasons like, a better competitor product or a bad experience with the product. Almost every day there are either product launches or new addition of variants to new products. Coming up with a new variant or a new product requires a high number of iterative experiments on the combination of ingredients that go into making the product, processes and designs. One of the objectives is also carefully choosing the right number of variants to present to the customer.

We compare the work of majorly three researchers in the area of presenting choices to customers and identify the conflict in their theories. Upon looking at them individually they all appear to be totally correct and highly researched but collectively they produce a conflict which we will try to resolve.

Here are our contributions in this regard:

1. We formalize the studies to a separate field as choice theory in marketing.
2. We build upon very convincing and well tested yet contradicting studies.
3. We try to rectify the contradiction.
4. We provide a first-hand formula for identifying number of choices to present to the consumer along with, which choices to present to the user using machine learning recommender systems.

In the following section we present the underlying studies along with the researchers upon which we build our theory.

Figure1. [8]
A typical store where a customer is bombarded with choices. Customer has to make a choice every time they visit a store, for example while identifying one particular variant from one brand, validating the older using variant with new products in the market, identifying replacement for older variant of a brand. A lot of such decisions to make the choice happen sub consciously.

II. STUDIES AND INFERENCES

A. Study of Howard Markowitz by Malcolm Gladwell

This refers to research of Malcolm Gladwell [1] where he talks about the food industry and its earlier assumptions. Gladwell makes remarks to how Howard Moskowitz [2][3] changed the food industry in three important ways:

- People don’t know what they want. They come to know of it when given with choices. For example: Consumers are more likely to give wrong answer if directly asked how they want their coffee. There can be many reasons for this e.g. they may not be honest or creative or insecure or out of unaware cognitive reasons. We can’t always explain what we want deep down. We seem to be manipulated at times by our own opinions about ourselves and feel confident about it. Malcolm gives multiple examples to validate this point [4].

- Universal to variability is incorrect. That there is no one kind of tomato sauce to satisfy the tastes of a billion people. There is no single perfect way to make a dish. There has to be different perfect "recipes" (plural intended [4]) for different people.

- Horizontal segmentation. Conventional way to market the products was that people want something higher, more expensive, something to make people aspire to, even if it was only for mustard sauce. Howard's idea is "The way to make more people happier, is by making different types of mustard for different types of people". [4]

Gladwell's conclusion "In embracing diversity of human beings we will find a surer way to true happiness" [4]. By this we can infer more varied products for such diversified population will bring more happiness.

B. Paradox of Choice by Barry Swartz

Refers to Barry Swartz on "paradox of choice" [6]. Barry tries to relate freedom with choice. The freedom he talks about is general. Barry states its conventional practice if you maximize choice you maximize freedom. We do not infer freedom in general sense instead we infer it as "freedom of choice" since freedom exists beyond consumerism.

Barry states to prove that having more choices make one take no choice at all [6]. Or if forced to choose consumer will feel miserable. The reason is that consumers actually care about what choice they make [9]. Consumers don't want to make a wrong choice whether if it's for eternity (like choosing a life partner) or even for something as trivial as choosing their salad dressing.

Barry's remarks that "if a person is made to choose from more choices be it anything, he/she will be less satisfied; Less satisfied compared to the choice he/she have had to make when there were lesser items to choose from, because then he/she will not be complaining; That he/she could have had better results with a choice that he/she left out" [6]. Barry states 3 reasons for the unhappiness caused by choosing:

- Imagination of alternative choices that one could have had made, causes "regret". In earlier times when people had less choices they happened to be happier in the given choice. Because they had less alternatives and had to go with the limited options provided.

- Opportunity costs - We value things which we compare them to. Attractive features of alternatives that we reject. Missing the opportunity costs of parking space when you are in vacation.

- Escalation of expectations - Barry gives example of buying jeans from slim fit, easy fit, and exact fit. This makes a person choose better but makes him/her feel worse. "good results but less satisfaction".

- Guilt - If the pair of jeans you bought turns out to be a bad fit, who's responsible for it? You! Since you chose it. Just the thought that you could have had made a better choice makes you feel bad. You blame yourself, even if you chose good.

We come to see the conflict of happiness from the two works. One study claims that secret to consumer happiness is more choices, where as another talk claims its less choices. Both of them show this statistically and with lots of proven examples. Barry says that there is a 2% decrease for every 10 new choices for mutual funds. It's true because it puts more hard work to choose from more choices. For the same reason we don't pay attention to 1% or 1.5% brokerage charges on our mutual fund investments but negotiate for few dollars in the daily market because the calculation of it makes it more difficult to choose from. Some may say that more choices leading to more happiness is in the case of food industry only. But we would like to argue that's not the case, it is more like both the results are specification of the same general thing.

C. Decision making illusions by Dan Ariely

Final study we would like to consider is by Dan Ariely "are we in control of making our own decisions" [9].

Dan talks about decision making illusions. He starts off by stating that if our minds can be tricked by visual illusions how can we confidently say that we are not making bad decisions unknowingly every day, especially when most of our life is based on choosing. He gives an example of filling a form and how users get tricked into selecting the default option. Because it can be that not selecting anything or avoiding yourself from making a choice will lead to another choice which you may not want either. As Dan puts it we have an illusion of deciding rather than really taking a decision. And this happens because we care, it's difficult and it's complex to make even small decisions. In order to not make a mistake or if it becomes too much work, in the moment we don't know what to do. We just pick what was chosen for us. A question arises "does this happen with the subject experts as well?". Dan provides an example. Physicians decision about surgery of a patient changes when two new medicines were suggested rather than suggesting only one new medicine. Most likely it becomes too much work.

III. IRRATIONAL DECISION MAKING

We have mainly two types of irrational decision making. First as we explained, we just pick what was chosen for us. The second type can be explained from dating example. Out of two people on a dating site (both of them comparably look good) the one becomes more popular whose similar uglier version is introduced. This is constantly used by marketers when providing offers, in malls, in mutual funds by providing an extra option to compare with. This tells we are not so much in control of making our decisions (it's the person who creates choices for us). Dan Ariely states that our mental world has limitations just like our physical world. Understanding our cognitive limitations will help us design better stock market, healthcare, retirement plans and other important stuff.

How can we solve this conflict? Dan Ariely states introduction of an inferior option (useless option which nobody wants) helps people figure out what they wanted. This is good and bad because it can be used to make you like an option, since that option has been made to look more attractive by introducing an inferior choice at the same cost.

According to Malcolm more choices, diversification of products leads to more happiness. According to Barry choosing from surplus choices make us feel miserable. Whereas Dan states introducing another option which is increasing number of choices helps people figure out what they want.

IV. Quantization of Choices

Barry's concluding remarks are 'some choices are better than none but more choice is not better than some choice'. Here 'some choice' has to be defined by our welfare, past our welfare is more choice.

It's not proper quantization, note that welfare is not limited. We always want more, things can always be improved. Now the sense we believe that he talks about is the same- whatever fills the needs, the present and near future requirement but it doesn't address the aesthetic/luxury part of the products or the products which need to be a lot aesthetic functionally.

Consumers should be given choices when they know what they want. Things with specific functionality - this has to do with requirements, with capability of products, with needs, but when it comes to fashion, taste, food, perfection; we feel this is defined by a personal taste. In that case people may not know completely well what they want, they may need choices, clothes to try on to known how they look, which looks better, same goes with food many a time. But when it comes to a car, a bike, a cell phone, people want better in lesser prices, it's often about more features.

We can say that there are two types of properties in a product, it's functional value - based on the present/future need. The other one is the creative value which includes brand, looks, appearance, garnish, ambience, art, etc. One may find drastic differences in the price value of the two properties. A piece of art has maximum creative value where as a computer has maximum functional value.

A producer cannot simply provide a smaller number of choices or a greater number of choices to the user. It should be a carefully calibrated number going through lots of iterative experimentation and experience. We provide the following method before launching a product to identify the number of choices to present to the user and also which choices or combinations to present so that those result in effective sales of the products. We make use of survey data for this.

V. Options Selection

Finding the number of options to present to the consumer is like finding the number of clusters in the userbase. There might be an overlap of some users which buy more than one variant of the product, we are going to ignore such cases. By carefully identifying the number of mutually exclusive features in the product or the ingredients. Different ingredient quantities should also be treated as mutually exclusive. All the variants are present to a sample of consumers to get to a survey data. We argue that the mutually exclusive features of the products are representative of the features of the group of users which are going to use that variant. Once this inference is concluded this problem essentially becomes the problem of clustering the users (which cluster around some of the variants) with prime objective to find the number of clusters.

We take survey data where all possible combination ingredients in possible product(s). All variants are presented to a sample of users. For example, we come up with **n** combinations of product variant and **m** people for sample survey. The **m** people based on their experience with the samples are asked to rate the products on a scale of 1 to 10. People can also be asked to rate simply 0 or 1 where 1 can be given to only one product by one user. This way we have **n*m** user product survey ratings.

Industries tend to work with statistical methods where they try to identify different peaks in the distributions. We instead look towards how machine learning approach can be used to identify such individual distributions. With the survey data we transform data to user-item interaction pair format. We assume **N** number of features for every variant and every reviewer. We split data in train and test for accuracy validation and run factorization models [10] to learn the ratings provided by the reviewer.

In the process of learning the reviews, we also get to learn the **N** features of the products. We learn the distribution by learning the effective number cluster of the products based on those features and look at every cluster distribution separately. We implement hierarchical clustering to manually look at the formation of clusters and decide number of clusters based on granularity. We can then implement any clustering method like K-Means to get the actual clusters. We identify the nearest points pertaining to the centroid of clusters and take them as the choices to present to the user.

We suggest matrix factorization-based recommender system approach for following reasons:

1. Even though the products have their set of ingredients and ratio to features to create the product variant which can be account as the variant features they are based on the feasibility of what is possible to create that variant. They are not based around user liking or purchase behavior.
2. The features learnt with factorization model are around people reaction to the product where they may or may not represent the physical features or combination of physical features of the products.
3. The clustering done on the factorized features is very much likely to give the proper clusters of user behavior around the product.
4. The centroid of those clusters or the products nearest to the centroid since the centroid may or may not represent the exact variant of the product. In this manner we can get to know which products gather maximum like behavior or each type of user and is also representative of similar product variants. In identifying so, we can discard all other variants in the cluster and consider the centroid (nearest to centroid) product candidates only for production.

Alternative to hierarchical clustering we list three other approaches which could have been used to identify the number of clusters.

A. Silhouette method

The average silhouette method [11] looks for quality of clusters and discards bad clusters which do not have good clustering as a group.

Briefly, it measures the quality of a clustering. That is, it determines how well each object lies within its cluster. A high average silhouette width indicates a good clustering.

Average silhouette method computes the average silhouette of observations for different values of k. The optimal number of clusters k is the one that maximize the average silhouette over a range of possible values for k.

The algorithm is similar to the elbow method and can be computed as follows:

1. Compute clustering algorithm (e.g., k-means clustering) for different values of k. For instance, by varying k from 1 to 10 clusters.
2. For each k, calculate the average silhouette of observations.
3. Plot the curve of average silhouette according to the number of clusters k.
4. The location of the maximum is considered as the appropriate number of clusters.

B. Dunn Method

The Dunn index [12] is another internal clustering validation measure which can be computed as follow:

1. For each cluster, compute the distance between each of the objects in the cluster and the objects in the other clusters
2. Use the minimum of this pairwise distance as the inter-cluster separation
3. For each cluster, compute the distance between the objects in the same cluster.
4. Use the maximal intra-cluster distance (i.e. maximum diameter) as the intra-cluster compactness
5. Dunn Index (D) is given by:
   D=minimum separation/maximum diameter

If the data set contains compact and well-separated clusters, the diameter of the clusters is expected to be small and the distance between the clusters is expected to be large. Thus, Dunn index should be maximized.

C. Elbow Method

This is the oldest method for finding number of clusters in a dataset. Multiple values of K starting from 2 clusters to get the cost of training. With a greater number of K, the value of cost function drops but the rate of drops after a certain point creating an elbow when looked at plot. Hence this is considered as a visual method and the K at elbow is chosen to be most effective number of clusters [5][7].

CONCLUSION

More choices give consumers more freedom to choose for their welfare. They can choose from better options or better prices among other factors. Quality can be anything, a time saving option can also be a quality. According to Barry we don’t want too many choices, that makes us miserable after buying. According to him the key to happiness is less expectations. The consumer can implement it on their side, it’s like a personal tap. But how to know, how many choices are too many choices? We explained Dan’s introduction to how one useless choice can help people know what they want. It is crucial decision for the business on the number of choices to present to the user and which choices to present to the user.

We presented how machine learning, factorization models can help us identify better and early how many product variants and which variants to present to the user. We provided steps to conduct such operation and also the explanation as to why it is useful in that manner.

This is a work in progress(subject to availability of appropriate data) where the study tries to identify in general how much product diversification is efficient diversification. If consumers get blinded by number of choices it results in not buying anything at all which poses risk of reduction in sales.

The study emphasizes on some products can be designed better. We encourage product owners, researchers to conduct the experiments on such datasets in their respective businesses.

REFERENCES

[1] Gladwell, M., 2004. The ketchup conundrum. New Yorker, 6.

[2] Moskowitz, H.R., 1970. Ratio scales of sugar sweetness. *Perception & Psychophysics*, *7*(5), pp.315-320.

[3] Moskowitz, H.R., Kluter, R.A., Westerling, J. and Jacobs, H.L., 1974. Sugar sweetness and pleasantness: Evidence for different psychological laws. Science, 184(4136), pp.583-585.

[4] Gladwell, M., Choice, happiness and spaghetti sauce,TED, 2004

[5] Kodinariya, Trupti M and Makwana, Prashant R, “Review on determining number of Cluster in K-Means Clustering”, vol. 1, no. 6,International Journal, 2013.

[6] Barry Schwartz, The paradox of choice, TED, 2005.

[7] Andrew Ng, Clustering with the K-Means Algorithm, Machine Learning, 2012.

[8] Photo by Peter Bond on Unsplash

[9] Dan Ariely, Are we in control of our own decisions, EG 2008

[10] Kula, M., 2015. Metadata embeddings for user and item cold-start recommendations. arXiv preprint arXiv:1507.08439.

[11] Rousseeuw, P.J., 1987. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. Journal of computational and applied mathematics, 20, pp.53-65.

[12] Ray, S. and Turi, R.H., 1999, December. Determination of number of clusters in k-means clustering and application in colour image segmentation. In Proceedings of the 4th international conference on advances in pattern recognition and digital techniques (pp. 137-143).